

# 対話的意思決定支援システムのための 視線運動を用いた興味変化推定

下西 慶†

川嶋 宏彰†

石川 恵理奈†

松山 隆司†

† 京都大学 大学院情報学研究科

本資料は第30回人工知能学会全国大会, 204-OS-23a-2 “対話的意思決定支援システムのための視線運動を用いた興味変化推定”の著者バージョンです。

## abstract

To choose one candidate from many alternatives on a catalog content is a fundamental part of our daily life. However, the goal is sometimes not defined at first and may be changed during decision process. Therefore, in order to build interactive decision support systems, temporal changes of users' interests need to be traced. We have proposed a method to represent users' interest as a mixture of aspects of items and to learn those aspects in a data-driven fashion. We extend this method by using proposed multinomial statistical test, which detects users' distinctive browsing behavior for short time multiscale window size. By learning those aspects not from whole gaze behaviors but from the distinctive periods detected multinomial exact test, we expect to obtain aspects that are well related to users' comparison behavior and to trace temporal changes of users' interests. We demonstrate the applicability of the proposed methods through experiments.

## 1. はじめに

複数候補の中から1つの対象を選択するような選択行動は我々が日常的に行っている行動である。そのような選択行動において、我々は常に明確な基準を持っているとは限らず、時として自らの選択基準が漠然とした状況から選択行動を開始し、複数の基準に基づいて比較を行うなどして、自らの選択基準を動的に変化させながら絞り込んでいくことがある [1]。このような場合、たとえば情報を獲得する際にもどのようにして情報の検索を行えばよいかわからない、また初期の検索結果は必ずしも自らの目的に沿ったものとなっているとは限らない、という問題が存在する。そのため、対話的にユーザの選択行動を支援することで解決をはかるシステムが提案されている [2, 3]。たとえば [2] では、ユーザに対してある候補とその候補の持つ特徴、およびその候補に対するいくつかの批評（もっと安く、など）を提示し、ユーザからの批評を反映して新たな候補の提示を繰り返すといった手法をとっている。

これまでに提案されてきている、対話的にユーザの選択行動を支援するようなシステムの多くは、ユーザの興味は潜在的には定まっていて変化しないことを仮定している。しかしながら、ユーザは選択行動において自らの興味を動的に変化させることが考えられる。そこで本研究においては、ユーザの興味のある空間上での分布として捉え、その興味の分布を収束させることによってユー

ザが自身の興味（選択基準）を明確にし、さらにその明確になった基準に基づいた選択を行うというプロセスの支援を目指す。

ユーザの視線運動はユーザの意思決定のさまざまな側面を理解するために重要な要素となることが知られている [4]。さらに、視線はユーザの内的な状態を反映してふるまうものであり、音声情報などと異なりユーザの明示的な行為を必要としない。本研究ではユーザが明示的に入力を与えることが困難であるような状況を想定しているため、ユーザの視線運動を興味を推定するための手がかりとする。

しかしながら、ユーザが同一の属性値を有するような対象を閲覧していたとしても、着目している基準に応じてその意味合いが異なることや、ユーザの興味の変化のパターンは一定ではないということが考えられる。すなわち、“どのような空間でユーザの興味を表現すべきか”、“どのようにしてユーザの興味の変化をとらえるべきか”という2つの問題を解決する必要がある。

興味の表現の問題に対しては、著者らの提案する手法として、ユーザが“なぜ見ているのか”という対象の持つ側面（アスペクト、たとえば、健康によさそう、など）に着目し、ユーザの興味をこのアスペクトの空間上で表現する手法（1）がある [5]。この興味空間の表現は、対象群の持つ各属性値はそれぞれ単一では多義的な意味合いを持つため、複数の属性値の組を考慮することで初めて興味の表現に用いることが可能なのではないかと考えている。

一方で興味の変化を捉える手がかりとしては、“短時間窓幅を用いてマルチスケールにユーザの閲覧行動からその偏りを検出する”という手法（2）を提案している [6]。この手法では、ユーザがある特定の属性値に着目した閲覧を行っている区間を検出することが可能である。

本研究では、ユーザの興味に変化する様子を捉えることが1つの課題となっているが、（1）の手法は、閲覧行動中のユーザの興味を一定であると仮定している。そこで本稿では、（2）の手法においてユーザの閲覧行動の変化が捉えられることを利用し、（1）の手法を拡張して興味に変化する状況に適用することを目指す。すなわち、ユーザが特定の属性値に着目しているような区間においては、その興味が一定であると仮定し、（2）の手法によってまずユーザが特定の属性値に着目している区間を検出し、その結果に応じてユーザの閲覧行動を分割する。ついで、その分割された区間に対しては興味が一定であると仮定して（1）の手法によって興味空間の学習、興味の推定を行うという手法を提案する（図1）。

## 2. アスペクトによる興味の表現法

この節では、“なぜ見ているのか”に着目したユーザの興味の表現法としてアスペクトおよび視線運動からアスペクトを学習する手法について述べる。

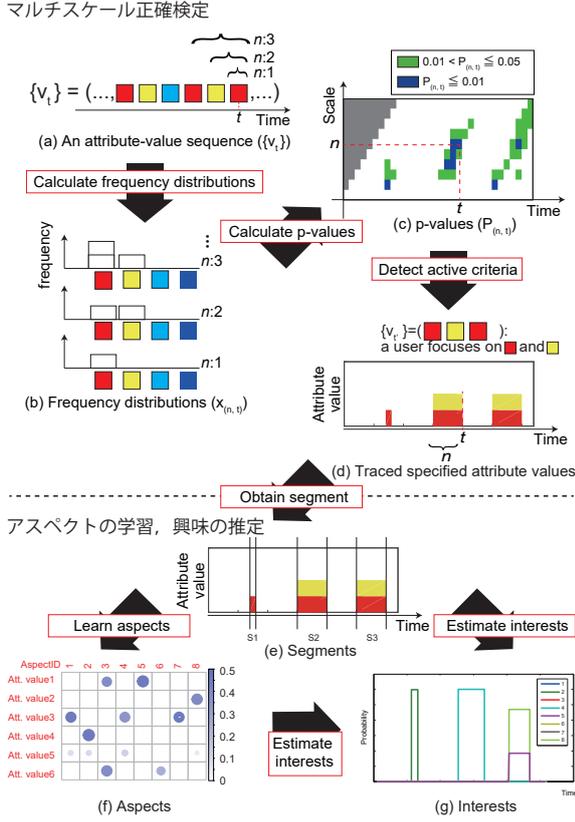


Figure 1: 提案する興味変化の推定法の概略

## 2.1 アスペクトによる興味の表現

ユーザがある対象を閲覧している場合、ユーザはその対象自身ではなく、その対象が持っている側面に着目していることが考えられる（例：健康によさそうなものだから見ている）。さらに、同じ属性値を持つ対象を閲覧していたとしても、その対象のドメインであったり従事しているタスクによって、その属性値の持つ意味合いは異なるものとなることが考えられる。そこで、各対象や属性値自身ではなく、複数の対象が共通して持つ側面（アスペクト）をユーザの興味の表現に利用し、これら複数のアスペクトに対する重視度という形でユーザの興味を表現する。すなわち、あるコンテンツに対して  $M$  種のアスペクト  $Z_1, \dots, Z_M$  が存在するとしたときに、ユーザの興味を、あるアスペクト  $Z_m$  への重視パラメタ  $\theta_m$  を用いて  $\theta = \{\theta_1, \dots, \theta_M\}$  ( $\theta_m \geq 0, \sum_m \theta_m = 1$ ) として表現する。さらに、これらのアスペクトは対象群が有している複数の属性値によって表現できると考えられるため、複数の属性値との関連度として特徴づける。

## 2.2 視線運動

ユーザがある基準に基づいて選択行動を行う際、ユーザはその基準を満たす複数の対象間で視線が遷移し、比較を行うのではないかと考えられる。そのため、本研究においては、ユーザの閲覧行動をそのサンプリング時刻ではなく閲覧対象の切り替わりタイミングに基づいて考える。すると、ディスプレイ上に提示されている対象集合  $\mathcal{I}$  に対して、ユーザの視線情報はその注視対象系列  $i_{1:T} = (i_1, \dots, i_T)$  ( $i_t \in \mathcal{I}, i_{t-1} \neq i_t$ ) として与えられる。たとえば図3に示すように、各対象は複数の共通した属性とその値を有している。ここでは  $J$  種の属性が存

在するとし、それぞれの属性  $j$  は  $K_j$  種の属性値  $\mathcal{V}^{(j)} = \{V_1^{(j)}, \dots, V_{K_j}^{(j)}\}$  をとるとする。これにより、ユーザの視線運動は注視対象系列からそれぞれの属性に対応した注視属性値系列  $v_{1:T}^{(j)} = (v_1^{(j)}, \dots, v_T^{(j)})$  ( $v_t^{(j)} \in \mathcal{V}^{(j)}, j = 1, \dots, J$ ) に変換可能である。

## 2.3 アスペクトの学習法

まず、各アスペクトを複数属性値との関連度として特徴づける、すなわち、あるアスペクト  $Z_m$  を、それぞれの属性値との関連度  $A_m^{(j,k)}$  を用いて、 $\mathbf{A}_m = \{A_m^{(1,1)}, \dots, A_m^{(J,K_J)}\}$  ( $A_m^{(j,k)} \geq 0, \sum_j \sum_k A_m^{(j,k)} = 1$ ) として表現する。ここで、ユーザの興味が一定である時間区間をセグメントと呼び  $s$  と表記する。このセグメント  $s$  においてユーザの興味が  $\theta_s = \{\theta_{s1}, \dots, \theta_{sM}\}$  ( $\theta_{sm} \geq 0, \sum_m \theta_{sm} = 1$ ) として与えられている場合、ある属性値  $V_k^{(j)}$  に対する重視度は、各アスペクトへの重視度とそれぞれの属性値との関連度の積和として、 $\sum_m A_m^{(j,k)} \theta_{sm}$  となる。

ここで、それぞれで興味が一定であるようなセグメント  $s$  の集合である  $S$  およびその閲覧行動として注視属性値系列から獲得された属性値集合  $\{v_{1:T_s}^{(1)}, \dots, v_{1:T_s}^{(J)}\}$  が与えられているとすると、推定すべきパラメタは、すべてのセグメント  $s \in S$  に対して共通するアスペクト集合  $\{\mathbf{A}_1, \dots, \mathbf{A}_M\}$  および各セグメントにおける興味の集合  $\{\theta_1, \dots, \theta_S\}$  である。

これらのパラメタは、EM アルゴリズムを用いて対数尤度

$$L = \sum_{s \in S} \sum_{t=1}^{T_s} \sum_{j=1}^J \sum_{k=1}^{K_j} \ln \left( \sum_m A_m^{(j,k)} \theta_{sm} \right) \quad (1)$$

を最大化することで求める。このモデルは観測データからそれらのデータをよく表現するような少数の軸を学習する手法の1つである確率的潜在意味解析 (probabilistic Latent Semantic Analysis; pLSA [7]) をベースにしたものである。

## 3. マルチスケール正確検定による着目属性値の変化推定

この節では、ユーザの閲覧行動からその偏りをマルチスケールに検出することでユーザが特定の属性値に着目しているような区間を捉える手法について述べる。

### 3.1 マルチスケール正確検定

提案するマルチスケール正確検定は、ユーザがどの属性値にも着目していない状況における閲覧行動をあらかじめモデル化しておくことで、ユーザの実際の閲覧行動から偏りを持った区間を検出する手法である。この検出においては、変化を捉えるために短時間窓を考慮して、少数サンプルに対する検定問題である正確検定を用い、さらにその窓の大きさをマルチスケールで考える。そのうえで検出された区間における閲覧対象を解析することで、その区間においてどのような属性値に着目していたのかを推定するというアプローチである（図1上部）。

まず、ユーザの閲覧行動の偏りを正確検定を用いて検出するために、ユーザがどの属性値にも着目していない閲覧行動をモデル化する。簡単のため、ユーザは特定の属性値に着目していない状態においてはコンテンツを一様に閲覧することを仮定する。これによって、ある属性値  $V_k$  の閲覧されやすさを表すパラメタ  $p_k$  は、コンテンツ内の

対象数  $N$ , コンテンツ内で属性値  $V_k$  を持つ対象数  $N_k$  を用いて,  $p_k = N_k/N$  として与えられる. これらのパラメータ  $\mathbf{p} = (p_1, \dots, p_K)$  の下で, 注視属性値度数分布  $\mathbf{x}_{(n,t)}$  の発生確率は, 多項分布  $f(\mathbf{x}; n, \mathbf{p}) = n! \prod_{k=1}^K p_k^{x_k} / x_k!$  に従うと考えられる.

ここで, ユーザが閲覧行動中にある属性値を持つ対象間を見比べているような区間を検出するため, 2.2 節と同様にして, 注視対象の切り替わりタイミングに基づいてユーザの閲覧行動から注視属性系列  $v_{1:T}$  を獲得する. そして, ある時刻  $t$  に対してマルチスケールの短時間窓を考慮し, そのサイズ  $n$  の窓内での部分注視属性値系列  $v_{t-n+1:t}$  ( $1 \leq n \leq 10$ ) における各属性値の度数分布  $\mathbf{x}_{(n,t)} = (x_1, \dots, x_K)$  ( $\sum_k x_k = n$ ) を計算する (図 1 (a) から図 1 (b)).

これらの度数分布  $\mathbf{x}_{(n,t)}$  に対する  $p$  値 ( $P_{(n,t)}$ ) は, 以下の式 (2) で計算される.

$$P_{(n,t)} = \sum_{\hat{\mathbf{x}}: f(\hat{\mathbf{x}}; n, \mathbf{p}) \leq f(\mathbf{x}_{(n,t)}; n, \mathbf{p})} f(\hat{\mathbf{x}}; n, \mathbf{p}). \quad (2)$$

マルチスケールに  $p$  値を計算することで, 図 1 (c) に示すように, 「いつ, どの程度の長さで」ユーザが偏りを持った閲覧行動を行っているかが検出できる.

ついで, この検出された区間において, ユーザがどの属性値に着目していたのかを調べるために, 元の注視属性値系列に立ち戻る. 1%水準で有意に差があるとされた区間において, ある属性値  $V_k$  に対する相対的な度数分布  $x_k/n$  が多項分布のパラメータ  $p_k$  より大きい場合にユーザがその属性値に着目しているとする (図 1 (d)).

#### 4. マルチスケール正確検定を用いた閲覧行動の分割に基づくアスペクト空間の学習および興味変化の推定

この節では, ここまでに述べてきた手法を用いて, どのようにしてユーザの興味の変化を推定するかということについて述べる.

1. 節で述べたように, アスペクトの学習, およびアスペクトを用いた興味の推定に対しては, 興味が一定であるような区間を定める必要があった. また, マルチスケール正確検定においては, 各属性ごとに特定の属性値に着目しているかどうかを検出するにとどまる. そこで, マルチスケール正確検定によってユーザが特定の属性値に着目していると検出された各区間においてはユーザの興味は区分的に一定であると仮定する. これにより, 2.3 節で述べたアスペクトの学習を, 区分的に興味に変化する状況に拡張することができる.

具体的には, 図 2 に示すように, マルチスケール正確検定でユーザが特定の属性値に着目していると推定された時区間において, ユーザが着目する属性値の集合に変化が観測された箇所ではユーザの閲覧行動を分割し, セグメント集合  $S$  を獲得する. そして, これらのセグメント  $s \in S$  に対するユーザの興味を  $\theta_s$  として, アスペクトの学習および興味変化の推定を行う.

#### 5. 実験

提案手法について, その着目属性値の検出に基づいたアスペクトの学習, およびその興味の変化が捉えられるかを実験参加者の協力により得られたコンテンツ閲覧時の視線データを用いて確認した.

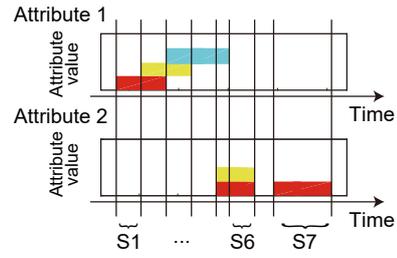


Figure 2: マルチスケール正確検定に基づくセグメントの獲得例



Figure 3: 実験環境および提示コンテンツの例

#### 5.1 実験設定

実験参加者には, 図 3 に示すような実験環境で 16 個の対象が提示されているコンテンツの閲覧を行ってもらった. さらに, 与えた教示に応じた対象を 1 つ選択するというタスクを行ってもらった際の視線運動を計測した. 実験参加者は 8 名であり, それぞれ 8 種のカタログコンテンツを閲覧してもらった. すなわち視線データ数は計  $8 \times 8 = 64$  (セッション) である. それぞれの対象は 4 つの共通した属性: カテゴリ, 価格, ランキング, レビュー, を持っている. またこれらの属性に対してはレビュー以外は 4 種, レビューは 5 種の値を取る.

#### 5.2 実験結果及び考察

計算された  $p$  値及び, 偏りがあると検出された区間においてユーザが着目していると推定された属性値の例をそれぞれ図 4 (a), 図 4 (b) に示す. このセッションにおいて与えた教示は, 「アルコール (4 番目のカテゴリ) の中で 4 つ星以上のレビューがついているものを選択してください」である.

さらにユーザの閲覧行動に対して, 各セッションで興味が一一定であるとしてアスペクトを学習した例, およびマルチスケール正確検定による推定結果を用いて閲覧行動からセグメントを獲得しアスペクトを学習した例を図 5 に示す. ここでは, 与えるアスペクトの数を, タスクの数と同数の 8 個として学習を行った.

図 4 (b) のように, ユーザはその閲覧行動の間に着目する属性値を変化させているため, セッション全体での興味を一定としてアスペクトを学習した場合は図 5 (a) のように, 大多数の属性値が 1 つのアスペクトに関連して学習されてしまっている. その一方で, 図 5 (b) を見ると, 少数の属性値が強い関連度を持って学習されており, さらに, 与えたタスクの 1 つである, 「4 番目のカテ

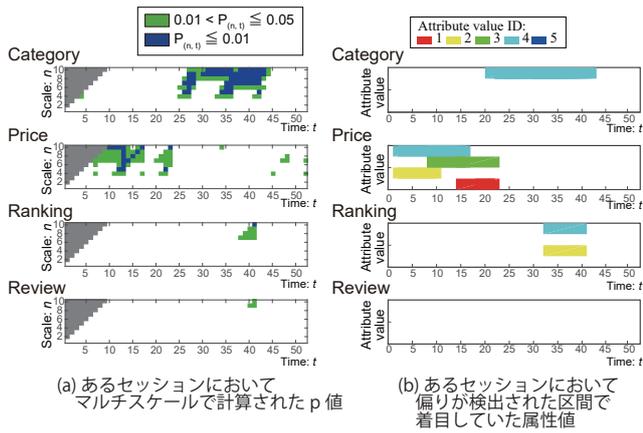


Figure 4: 計算された  $p$  値, および着目していると推定された属性値の変化の例. 各行はそれぞれの属性 (カテゴリ, 価格, ランキング, レビュー) を示す. 時刻  $t$  は注視対象の切り替わりに基づいている.

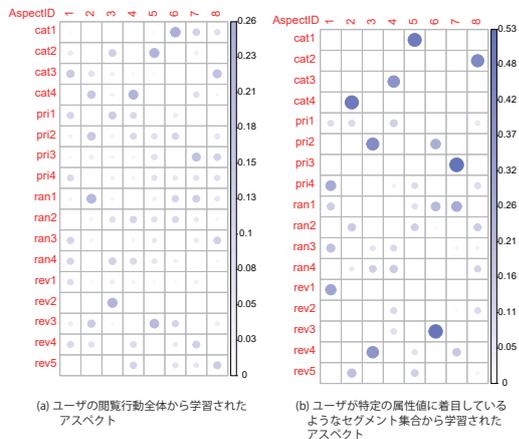


Figure 5: 学習されたアスペクトの例, アスペクト数は 8 であり, 各列がそれぞれのアスペクトと属性値との関連度を表す.

ゴリでありレビューが 5 のもの」が 2 番目のアスペクトとして学習されていることなどから, ユーザが比較行動を行う際に特に着目されやすい属性値との関連度が学習されていると考えられる. 図 6 は, 図 4 のセッションに対して, このアスペクトを用いてユーザが特定の基準に着目した閲覧行動を行っている際の興味の推定結果の例である. 図 4 (b) を見ると, ユーザは  $t \geq 20$  ではタスクで与えられている属性値である 4 番目のカテゴリを持つ対象に着目した閲覧を長期的に行っていることが分かるが, 図 6 から, そのような区間においてもユーザの興味が変化している様子がマルチスケールで捉えられている.

## 6. まとめと今後の展望

本稿では, ユーザが曖昧な興味に基づいて対象を閲覧する際に, その興味を閲覧対象の持つ側面 (アスペクト) を用いて表現し, その時間的変化を捉えるための手法について述べた.

この手法においては, 個々の属性値をカテゴリカルな属性値として独立に扱っているが, たとえば価格などの

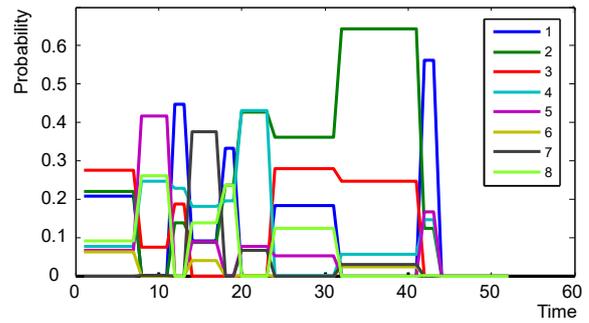


Figure 6: 推定された興味変化の例. 時刻  $t$  は注視対象の切り替わりに基づいている.

ように順序性を持つ属性値のようにある閾値のみに興味がある場合であったり, 個々の属性値単体では重要ではなく, 複数の属性値を同時に満たすことが重要となる場合も考えられる. また, ユーザの興味の变化は捉えられているが曖昧さの表現などはできていない. そのため, 順序性を持つ属性値の扱いや, 複数の属性値を同時に扱うための手法およびユーザの興味の曖昧さの表現について検討を行う. さらに, 実際に推定された興味に基づいてコンテンツを切り替える, 音声による追加の情報提示を行うなど, ユーザに対して働きかけを行い, それに対するユーザの応答 (視線運動の変化など) を観測することによって, ユーザの興味の推定を補正しつつその興味を収束させていく, といった対話的に意思決定を支援するシステムを構築する.

## 謝辞

本研究の一部は科研費研究補助金 15J06965, 13J05396, 26280075 および JST さきがけの補助を受けて行った.

## References

- [1] Sudatta Chowdhury, Forbes Gibb, and Monica Landoni. Uncertainty in information seeking and retrieval: A study in an academic environment. *Information Processing & Management*, 47(2):157 – 175, 2011.
- [2] Robin Burke. Interactive critiquing for catalog navigation in e-commerce. *Artif. Intell. Rev.*, 18(3-4):245–267, December 2002.
- [3] Joseph Polifroni and Marilyn Anne Walker. Intensional summaries as cooperative responses in dialogue: Automation and evaluation. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL)*, Columbus, June 2008.
- [4] Jacob L. Orquin and Simone Mueller Loose. Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1):190 – 206, 2013.
- [5] 下西 慶, 石川 恵理奈, 米谷 竜, 川嶋 宏彰, and 松山 隆司. 視線運動解析による興味アスペクトの推定. *ヒューマンインタフェース学会論文誌*, 16(2):103–114, 2014.
- [6] Kei Shimonishi, Hiroaki Kawashima, Erina Schaffer, and Takashi Matsuyama. Tracing temporal changes of selection criteria from gaze information. In *Companion Publication of the 21st International Conference on Intelligent User Interfaces, IUI '16 Companion*, pages 9–12, New York, NY, USA, 2016. ACM.
- [7] Thomas Hofmann. Probabilistic latent semantic analysis. *Uncertainty in Artificial Intelligence*, pages 289–296, 1999.